IEEE Transactions on Consumer Electronics, Vol. 43, No. 4, NOVEMBER 1997

AN EFFICIENT LAPPED ORTHOGONAL TRANSFORM IMAGE CODING TECHNIQUE

Sunil Kumar, Yogesh N. Rathi and R. C. Jain
Electrical and Electronics Engineering Department
Birla Institute of Technology & Science
Pilani 333 031 (India)

ABSTRACT

A fast and computationally less complex coding technique is described which uses partial-LOT computation algorithm and efficiently discards perceptually insignificant high frequency transform coefficients. The coding process involves AC energy classification, human visual system weighted normalization and quantization. The values of normalization factors are image independent and governed only by the bit-rate and activity index of the image blocks. Zones of various shapes and sizes are chosen for each activity class to perform zonal sampling which efficiently discards high frequency coefficients having zero values. A progressive transmission version of the proposed technique is also given. This technique gives better performance at comparable bit-rate than that obtained by using LOT/CVQ and LOT/VQ schemes.

I- INTRODUCTION

Image compression is playing a very important role in the areas of video conferencing, video phone and other consumer entertainment applications. Amongst the numerous techniques, transform coding is one of the most efficient methods of compressing images. However, one serious drawback of traditional transform coding, such as discrete cosine transform (DCT), at low bit-rates is the visibility of block structure, known as 'blocking effect', which is a natural consequence of the independent processing of each block. This drawback is overcome in a new class of overlapping block transforms [1-4]. Lapped Orthogonal Transform (LOT) is one such real, separable and fast overlapping block transform whose basis functions extend beyond the traditional image block boundaries. For example, in LOT of size N, each block has L samples, with L > N, so that neighboring blocks overlap by (L-N) samples. Because LOT maps L samples of each block into N transform coefficients, there is no increase in the bit rate compared to traditional transforms. To obtain the original input signal by the inverse LOT, subsequent output blocks must be overlap added. The properties of LOT along with a fast computable LOT matrix are given in detail in [1-4].

Rao et al. [5-6] have proposed two adaptive LOT coding schemes using classified vector quantization. However, they are computationally complex because each block of LOT coefficients is divided into several sub-vectors which are then vector quantized. A fast and variable bit-rate coding scheme is presented in this paper, whose coding complexity and bit-rates are lower than other adaptive LOT coding techniques discussed in the literature. In LOT, amplitude and variances of high frequency coefficients, lying beyond a 6x6 zone corresponding to first six rows and six columns of 8x8 coefficient matrix, become negligibly small after human visual system (HVS) weighting [5]. Therefore, we have used a significantly less complex LOT computation algorithm to losslessly compute only the 36 coefficients lying within the 6x6 zone. Then each sub-block, of 6x6 LOT coefficients, is classified into one of the four activity classes according to its directional activity. In order to further reduce the computational complexity, HVS weighted normalization and quantization is performed only on those coefficients (< 36) of a block, which lie within the largest zone corresponding to their activity class. The normalization factors are image independent and have been chosen empirically after testing on several images. After this, several zones of different shape and sizes, chosen according

Contributed Paper

Manuscript received April 28, 1997

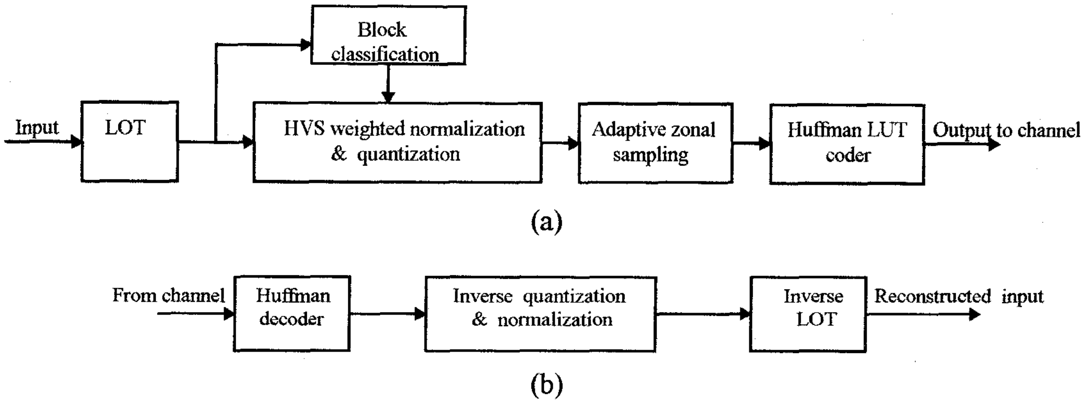


Fig 1 : Block diagram of the proposed LOT scheme.
(a) Coder. (b) Decoder.

to amount of AC energy present in the blocks of an activity class, are used to discard most of the zero valued quantized coefficients without further increasing the error. After zonal sampling, the coefficients are arranged in zig-zag order and Huffman coded. All the above mentioned features, without degrading image quality, significantly reduce computational complexity due to reduction in the number of addition and multiplication operations and also increase the compression ratio by reducing the number as well as amplitude of the coefficients to be coded. Moreover, the computational complexity of decoder is also decreased considerably as only less than 36 coefficients are required to be processed as compared to 64 coefficients in other adaptive techniques reported in the literature. This technique is also amenable to the progressive transmission of data in several stages until the final desired image quality is achieved, which is useful in interactive visual communication such as browsing.

In the next section, system configuration is described which includes transformation, activity index calculation, normalization and quantization, adaptive zonal sampling, coding and progressive image transmission. Section III contains the results of computer simulation followed by the conclusion in Section IV.

II- SYSTEM CONFIGURATION

Block diagram of the proposed coding scheme is shown in Fig. 1. The following subsections describe the coding scheme in detail.

i) 2-D LOT Computation

Input image is first partitioned into 8x8 non-overlapping blocks, and then 16x16 overlapping blocks are obtained as shown in Fig.2. Since, direct computation of 2-D LOT and its inverse requires large computation time, a computationally efficient method, described below, is used.

$$F = P^t X P = Z^t P \qquad ...(1)$$

where, $Z = X^t P$ and Z^t is its transpose. X and F are 16x16 input pixel and 8x8 coefficient matrices respectively. P is 16x8 fast computable matrix containing basis functions defined in [3]. To

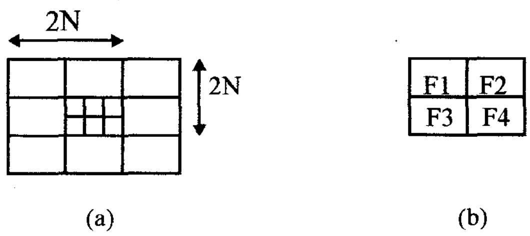


Fig 2 : (a) Overlapping of adjacent 2Nx2N pixel blocks
(b) Adjacent NxN LOT coefficient blocks

reconstruct the NxN shaded portion of the block shown in Fig. 2(a), inverse transformation is applied to each of the four adjacent LOT coefficient blocks (shown in Fig. 2(b)) as given below.

$$X_i = P F_i P^t , \quad i = 1,2,.,4 \qquad ... (2)$$

The overlapping common portion of reconstructed neighboring 2Nx2N blocks, X_1, X_2, X_3, and X_4 are then added. Since, the proposed coding scheme never uses a zone of size greater than 6x6, we have used a lossless partial-LOT computation algorithm in which first of all a 16x6 Z matrix is computed using (1), where order of X is unchanged while matrix P has been changed to 16x6. Then, a 6x6 LOT coefficient matrix F is computed using (1). Similar algorithm is used in inverse LOT computation. Use of this pruning algorithm results in substantial reduction in computational complexity due to reduction in addition and multiplication operations.

ii) AC Activity Index Calculation

The blocks of LOT coefficients are classified into four perceptual classes: one low activity class and three activity classes corresponding to the three edge orientations - horizontal, vertical and diagonal- based on three activity indices calculated using the energy of certain coefficients [5]. If all the three indices are below an empirical threshold (Th), the block is classified as a low-activity block. In rest of the discussion in this paper, these activity classes will be denoted by 1,2,3 and 4 respectively.

iii) Normalization and Quantization

Since HVS response to transform coefficient is dependent on its spectral range, significance of even a coefficient having high variance can be small to an observer if it corresponds to a spectral range for which HVS sensitivity is low. Hence each LOT coefficient is divided by normalization factor(D), chosen according to HVS weighting [7]. The values of normalization factors chosen by us (see Fig. 3) are image independent, as they have been optimized on several images of different characteristics, and are kept more for higher activity classes than those for low activity class to take advantage of distortion masking in subjective assessment of image quality. The quantizer simply converts a floating point to its nearest integer.

iv) Adaptive Zonal Sampling

Most images have a low pass energy spectrum and the high frequency (HF) LOT coefficients have low amplitude as well as low HVS sensitivity. Therefore, HF coefficients lying beyond a prespecified zone are discarded in zonal sampling. Since, the amplitude of HF coefficients in a block is indicated by it's activity class, zone size should be

8	2	3	5		
2	7	10			
3	10				
5					

Low

8	4	7	13	27	71
4	11	10	29	71	154
4	16	12	40	92	
10	23	48	52		
21	47	62			
47					

Horizontal

8	5	9	21	43	
5	7	19	47	94	
6	36	25	64	123	
10	29	64	104		
27	47	62			
71	77				

Vertical

8	5	4	10	21	35
5	9	10	29	71	154
4	10	12	32	46	194
8	23	40	75	69	182
27	47	92	95	120	
94	102	97	136		

Diagonal

(a)

8	4	7	18		
4	9	20			
7	20				
18					

Low

8	5	8	24	234	247
5	9	16	146	270	538
13	81	95	185	323	
134	133	185	297		
123	270	323			
247	538				

Horizontal

8	5	10	58	127	
5	9	65	128	269	
9	71	91	185	354	
18	128	177	297		
124	270				
129	589				

Vertical

8	5	8	32	118	259
5	9	10	128	259	563
7	10	21	177	336	709
26	122	169	271	500	645
118	247	321	500	800	
247	512	645	909		

Diagonal

(b)

Fig. 3: Values of normalization factors for different activity classes. (a) Lower values, D1 (b) Higher values, D2.

chosen depending on the activity class of the block. Moreover, energy distribution may vary even for the blocks of the same activity class. This suggests that the use of only one zone per activity class may either discard several significant HF coefficients in some blocks or retain some insignificant HF coefficients. Therefore, to keep reconstruction error within limits, several zones of different shape and sizes have been chosen for each activity class according to amplitudes of the coefficients and HVS sensitivity, after testing on several images, as shown in Fig.4. The smallest zone, beyond which all the quantized coefficients are zero, is finally assigned to each block.

v) Coding

After zonal sampling, the coefficients are arranged in zig-zag order [8]. Coefficient F(0,0) is coded using an 8 bit uniform quantiser. Similarly coefficients

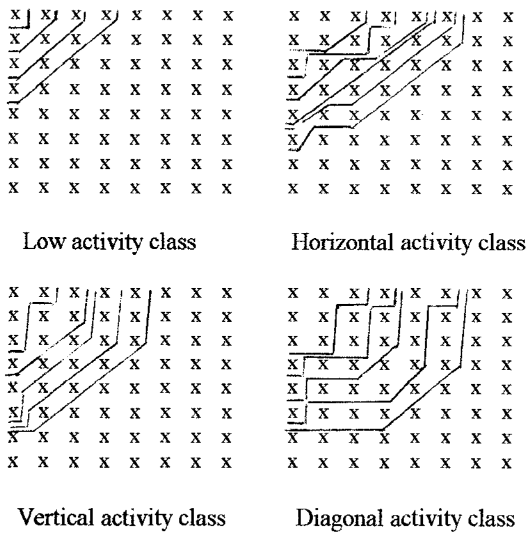


Fig. 4: Adaptive zones for blocks of different activity classes.

F(0,1) and F(1,0) belonging to activity classes 2, 3, and 4 are coded with 7 bits each. The remaining zero and non-zero coefficients are coded by runlength and amplitude lookup tables respectively. Both the lookup tables are truncated Huffman codes as used by Chen and Pratt [8] to further reduce the bit-rate. Amplitude lookup table also includes an 'end of block' (EOB) code to terminate coding of the block as soon as last significant coefficient of the block has been coded. Since, each zone used in the proposed coder contains a unique number of coefficients, the EOB code automatically identifies the class as well as zone to which a block belongs. This eliminates the need of overhead bits to encode class and zone information for a block.

vi) Progressive Image Transmission

Progressive image transmission (PIT) is becoming increasingly popular in applications such as browsing, requiring interactive image transmission, over narrow bandwidth channels. To construct recognizable image in PIT, minimum number of bits necessary to represent each block are transmitted first, which can then be progressively improved by transmitting additional bits in several stages until the desired quality is achieved. Because the coding scheme discussed in this section uses several zones for each activity class, it permits PIT by first transmitting the coefficients of only the smallest zone of an activity class to get the recognizable image followed by coefficients belonging to larger zones of each activity class.

III - SIMULATION RESULTS

The proposed coding scheme has been simulated on a PC. The original 512x512, 256 gray level test images 'Lena' and 'Palace', which have different characteristics and statistics, are shown in Fig.5.Fig. 6 shows the photographs of both the reconstructed test images, having excellent quality, using the proposed LOT coder at bit-rates of 0.70 and 0.79 bit/pixel (bpp) and SNR of 35.4 dB and 32.4 dB respectively. The subjective quality of reconstructed images was judged by displaying them at reduced spatial resolution. However, minor distortions in hair region and ringing near hat of 'Lena' are not clearly visible in photographs.The output bit-rate in the proposed coding scheme can be controlled by using the following three methods:

a) Choice of Threshold: More and more edge blocks get classified into low activity class when value of threshold (Th) is increased (SeeTable I). This reduces the bit-rate, as shown in Fig.7, because the number of coefficients in the largest zone of low activity class are less than most of the larger zones of edge classes. This, however, does not adversely affect the subjective image quality as only HF coefficients are discarded.

b) Choice of Normalization Factors: As value of normalization factor is increased, the amplitude of coefficients become smaller and HF coefficients tend to zero resulting in the use of smaller zones. For example, in Table I at a Th of 500, more blocks belong to smaller zones in each activity class after being normalized by D2 than using the lower values (D1), which results in lower bit-rate.

c) Elimination of HF LOT Coefficients: Successive elimination of LOT coefficients belonging to larger zones of each activity class, without affecting the encoding of coefficients in the remaining zones, reduces bit-rate. However, the image quality gets deteriorated when large number of HF coefficients are discarded. This method is, however, well suited for progressive image transmission.

Subjective quality of 'Lena' in Fig. 8(a) (at Th = 20000) is quite close to that in Fig. 6(a) (at Th = 500) which demonstrates its slow deterioration with increase in value of threshold (Method (a)). At the

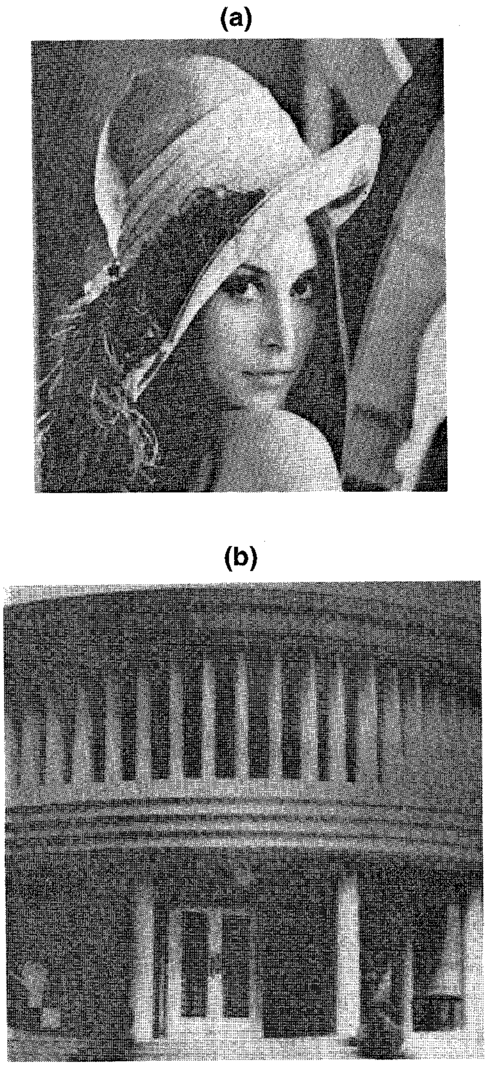


Fig. 5 : Original test images
(a) Lena (b) Palace

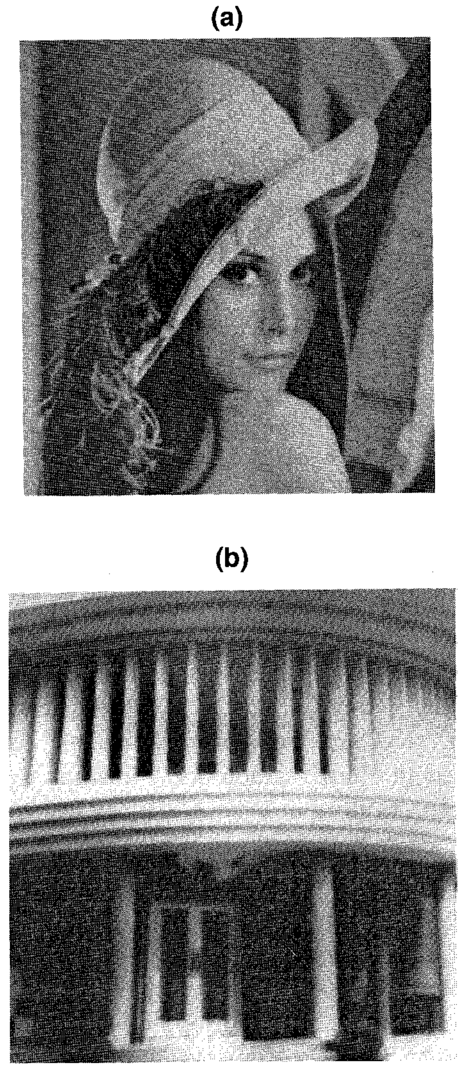


Fig. 6 : Coded images. (D1, Th = 500)
(a) Lena (b) Palace

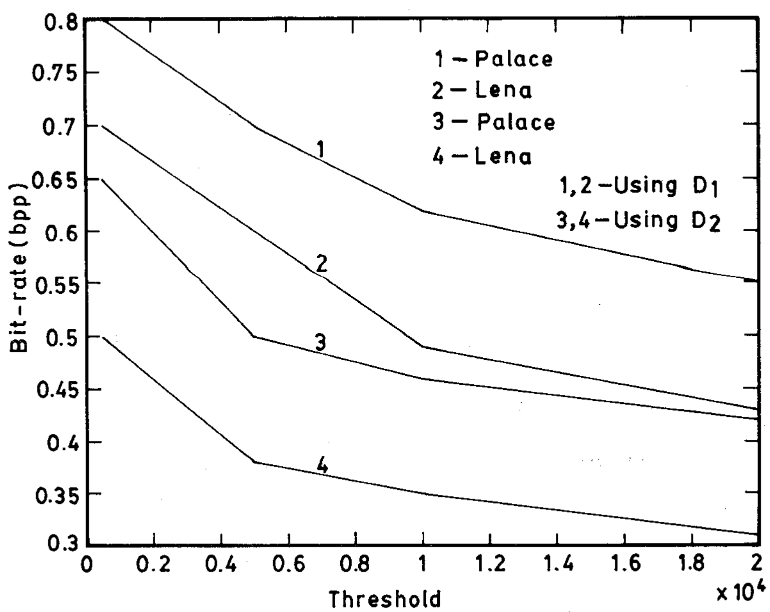


Fig. 7 : Threshold Vs Bit-rate curves for test images.

same bit-rate, subjective quality as well as SNR of 'Lena' in Fig. 8(a) is better than that in Fig 8(b). This is because, in Method (a) only few HF coefficients are discarded and the rest are unchanged whereas in Method (b) even the low and middle frequency coefficients are scaled down by a greater amount. Similarly, at the same bit-rate, quality of test image 'Lena' shown in Fig. 9(d) using the Method (c) (at Th = 5,000 and D1) is much better than that in Fig. 8(b) which has been obtained using Method (b). This is due to the fact that only HF coefficients which have low variance and poor HVS sensitivity are discarded in Method (c) while other coefficients are unchanged.

As shown in Table II, the proposed coding scheme gives SNR higher than or equal to that obtained by the LOT/VQ method described in [6]. 'Lena' coded at a bit-rate of 0.33 bpp, using the proposed coder is shown in Fig. 8(c), which has lesser ringing and better edge reproduction than in [6]. The PIT coding of images has been done at a target bit rate of 0.58 bpp and the reconstructed images are shown in Fig. 9. It can be clearly seen that subjective quality as well as SNR values (see Table III) of reconstructed image is superior than that reported in [5].

Table I : Number of Blocks in Each Zone at Different Thresholds

Zone	Threshold ($\times 10^3$)							
	0.5	2	5	10	15	20	25	0. 5*
10	1087	1766	2129	2383	2524	2603	2650	39
11	745	778	780	780	780	780	780	588
12	500	516	516	517	517	517	517	1112
13	93	95	95	95	95	95	95	686
20	8	8	7	4	3	3	1	0
21	43	39	34	22	14	8	0	0
22	174	147	106	64	41	26	12	4
23	344	217	140	86	50	30	13	35
24	370	127	52	22	14	9	5	688
25	50	6	2	1	1	0	0	263
30	0	0	0	0	0	0	0	0
31	32	29	21	11	4	1	0	1
32	46	36	20	8	3	1	0	3
33	171	90	38	19	7	2	1	237
34	17	5	2	0	0	0	0	25
40	9	9	8	6	3	2	0	0
41	43	42	33	20	9	4	3	0
42	140	116	77	32	16	5	3	3
43	203	66	31	14	7	4	1	353
44	20	0	0	0	0	0	0	59

- Zones 10 to 13 belong to low activity class and are in descending order of their size. Same is also true for other zones.

* Normalized by higher values (D2) while other columns are normalized by lower values (D1).

Table II : SNR Values for Lena using LOT/VQ and Proposed Coder

Bit-Rate(bpp)	LOT/VQ(dB)	Proposed Coder(dB)
0.33	30.6	32.5
0.50	33.7	34.0
0.70	35.5	35.4

Table III : Bit-Rates and SNR for LOT/HVS/CVQ and Proposed Coder at Each Progressive Stage

Bit-Rates (bpp)	SNR(dB) LOT/HVS/CVQ	SNR(dB) Proposed Coder
0.16	18.7	---
0.21	---	27.9
0.31	21.0	30.7
0.41	---	32.7
0.43	23.8	---
0.50	25.8	34.0
0.57	26.6	34.7

IV- CONCLUSION

A computationally efficient coding technique using LOT is discussed and compared with LOT/CVQ and LOT/VQ schemes. The simulation results show that the proposed technique not only requires relatively much less number of multiplication and addition operations and gives higher speed but also gives subjectively improved reconstructed images at equal or higher SNRs than those obtained by using LOT/CVQ and LOT/VQ schemes. The main features of the coding scheme are:

- For a 16x16 pixel block, only 6x6 coefficients are required to be computed as compared to 8x8 in other techniques. This is achieved by using a lossless partial-LOT computation algorithm which significantly reduces hardware complexity, which in turn reduces LOT computation time by about 30%.

(a)

(b)

(c)

Fig. 8 : Coded Lena image.
(a) Using D1 & Th = 20,000 (0.50 bpp & SNR = 34.0 dB).
(b) Using D2 & Th = 500 (0.53 bpp & SNR = 33.2 dB).
(c) Using D2 & Th = 20000 (0.33 bpp & SNR = 32.5 dB).

(a)

(b)

(c)

Fig. 9 : Progressive stages for Lena (D1, Th = 5000).
(a) First stage at 0.21 bpp. (b) Second stage at 0.33 bpp
(c)Third stage at 0.41 bpp. (d) Fourth stage at 0.50 bpp
(e) Fifth stage at 0.58 bpp

(d)

(e)

Fig. 9 : Continued

- Once the AC activity class of a block is known, further processing of coefficients in the coder is done only on those coefficients (< 36) which lie within the largest zone for that class. This further reduces the number of coefficients to be processed.
- The normalization factor values depend only on AC activity class of the blocks and do not change with the image characteristics.
- The use of adaptive zones reduces the runs of consecutive zeros in the coder, and the number of coefficients used in decoder. This reduces the computational complexity as well as bit-rate without affecting image quality.
- At a given bit-rate, impairment in quality is much less for Method (a) than for Method (b) given above.
- The reconstructed images rapidly converge to a good quality both subjectively and objectively when used in PIT mode.

REFERENCES

1. H. S. Malvar and D. H. Staelin, " The LOT: Transform coding without blocking effects," IEEE Trans. Acoust. Speech Signal Process., vol 37, pp 553-559, 1989.

2. P. M. Cassereau, D. H. Staelin and G. De Jager, " Encoding of images based on a lapped orthogonal transform," IEEE Trans. Commun., vol 37, pp 189-193, 1989.

3. H. S. Malvar, " Lapped transforms for efficient transform/ subband coding," IEEE Trans. Commun., vol 38, pp 1040-1044, 1990.

4. M. Temerinac and B. Edler, " A unified approach to lapped orthogonal transforms," IEEE Trans. Image Processing, vol 1, pp 111-116, 1992.

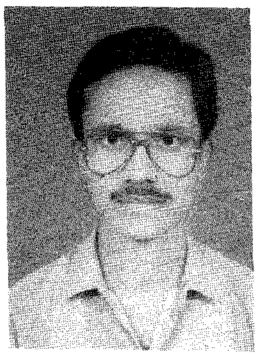
5. C. Hwang, S. Venkatraman and K. R. Rao, "Human visual system weighted progressive image transmission using lapped orthogonal transform/ classified vector quantisation," Opt. Eng., vol 32, pp 1524-1530, 1993.

6. S. Venkatraman, J. Y. Nam and K. R. Rao, "Image coding based on classified lapped transform - vector quantization," IEEE Trans. Circuits Syst. Video Technol., vol 5, pp 352-354, 1995.

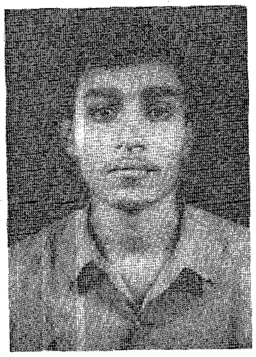
7. B. Chitprasert and K.R. Rao," Human visual weighted progressive image transmission," IEEE Trans. Commun., vol 38, pp 1040-1044, 1990.

8. W. H. Chen and W. K. Pratt, "Scene adaptive coder", IEEE Trans. Commun., vol 32, 1984, pp 225-232.

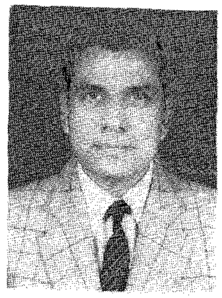
BIOGRAPHY



Sunil Kumar received his B. Engineering degree from Regional Engineering College, Surat and received M.E and Ph.D. in Electronics Engg. from Birla Institute of Technology & Science, Pilani in 1993 and 1997 respectively. He has been with Electrical and Electronics Engineering department of BITS, Pilani since 1990. His research interests include signal processing & image/ video compression.



Yogesh. N. Rathi received his M.Sc(Hons) in Mathematics and B.E(Hons) in Electrical & Electronics Engineering from Birla Institute of Technology and Science, Pilani in 1997. His research interests include image/video compression.



R. C. Jain received his B.E and M.E degrees in Electronics Engineering from University of Roorkee in 1970 & 1978 respectively. He also received his Ph.D. in Electrical Engineering from University of Alberta, Canada in 1988. He has been with the Electrical and Electronics Engineering Department of Birla Institute of Technology and Science, Pilani since 1989 where he is currently a Professor. His research interests include speech and video compression.